

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2015-520470

(P2015-520470A)

(43) 公表日 平成27年7月16日 (2015.7.16)

(51) Int. Cl. F I テーマコード (参考)  
**G 0 6 T 7/00 (2006.01)** G 0 6 T 7/00 3 5 0 Z 5 L 0 9 6

審査請求 未請求 予備審査請求 未請求 (全 32 頁)

(21) 出願番号 特願2015-518531 (P2015-518531)  
 (86) (22) 出願日 平成25年6月19日 (2013.6.19)  
 (85) 翻訳文提出日 平成26年12月15日 (2014.12.15)  
 (86) 国際出願番号 PCT/US2013/046447  
 (87) 国際公開番号 W02013/192253  
 (87) 国際公開日 平成25年12月27日 (2013.12.27)  
 (31) 優先権主張番号 13/530,925  
 (32) 優先日 平成24年6月22日 (2012.6.22)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 500046438  
 マイクロソフト コーポレーション  
 アメリカ合衆国 ワシントン州 9805  
 2-6399 レッドモンド ワン マイ  
 クロソフト ウェイ  
 (74) 代理人 100107766  
 弁理士 伊東 忠重  
 (74) 代理人 100070150  
 弁理士 伊東 忠彦  
 (74) 代理人 100091214  
 弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 データベース生成及び更新に深度ベースのトラッキングを用いた顔認識自己学習

## (57) 【要約】

顔認識トレーニングデータベース生成手法実施形態が提示され、概して、ある時間にわたって、及び人がある環境の中を進むときに、捕捉された人の顔の特徴を収集して、この人のための顔特徴のトレーニングデータベースを作成することを伴う。顔特徴がある時間にわたり捕捉されるので、この顔特徴は、様々な角度及び距離、種々の解像度、並びに種々の環境条件（例えば、照明及びヘイズ条件）下から見られたとおりの人の顔を表すことになる。さらに、人の顔特徴が周期的に収集される長期間の時間にわたり、これらの特徴は、人の見掛けにおける展開を表すことができる。これは、顔認識システムにおける使用のための豊富なトレーニングリソースを生み出す。さらに、人の顔の認識トレーニングデータベースが、顔認識システムによって必要とされる前に確立されることができるため、一旦採用されると、トレーニングがより迅速になることになる。

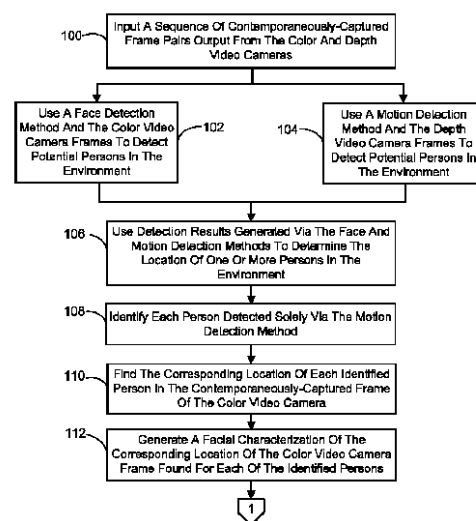


FIG. 1A

**【特許請求の範囲】****【請求項 1】**

環境に位置するとして検出された各人のための顔認識トレーニングデータベースを生成する、コンピュータにより実施されるプロセスであって、

コンピュータを用いて下記のプロセス動作を実行することを含み、該プロセス動作は、

(a) 同時捕捉されたフレームペアのシーケンスを入力するステップであり、各フレームペアはカラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む、ステップと、

(b) 顔検出方法と前記のカラービデオカメラフレームとを使用して環境において潜在的な人を検出するステップと、

(c) 動き検出方法と前記の深度ビデオカメラフレームとを使用して前記環境において潜在的な人を検出するステップと、

(d) 前記顔検出方法と前記動き検出方法とを介して生成された検出結果を使用して前記環境において 1 又は複数の人の位置を決定するステップであり、前記顔検出方法を介して生成された検出結果は、検出された各人について、人の顔を描写するカラービデオカメラフレームの部分の顔特徴を含む、ステップと、

(e) 単に前記動き検出方法を介して検出された各人について、

人の対応する位置を、前記カラービデオカメラの同時捕捉されたフレームにおいて識別するステップと、

前記人の顔を描写するカラービデオカメラフレームの部分の顔特徴を生成するステップと、

(f) 前記環境において検出された各人について、

人のために生成された各顔特徴を、前記人のために確立された不明人識別子に割り当てるステップと、

前記顔特徴の各々を、前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人の同一性を確定することを試みるステップと、

前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

を含む、プロセス。

**【請求項 2】**

同時捕捉されたフレームペアの新しいシーケンスを入力するステップであり、各フレームペアは、カラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む、ステップと、

プロセス動作 (b) 乃至 (e) を繰り返すステップと、

前記環境において検出され、同時捕捉されたフレームペアの前記新しいシーケンスにおいて描写された各人について、

人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応するかを判定するステップと、

前記人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応すると判定されるたび、前記人の同一性が事前に確定されたかを判定するステップと、

前記人の同一性が事前に確定されたと判定されるたび、同時捕捉されたフレームペアの前記新しいシーケンスから生成された各顔特徴について、該顔特徴が、前記人のために確立された顔認識トレーニングデータベースに割り当てられた各顔特徴から所定程度まで異なるかを判定するステップと、

同時捕捉されたフレームペアの前記新しいシーケンスから生成された各顔特徴について、該顔特徴が、前記人のために確立された顔認識トレーニングデータベースに割り当てられた各顔特徴から所定程度まで異なると判定されるたび、前記顔特徴を前記人の

10

20

30

40

50

ために確立された顔認識トレーニングデータベースに割り当て、前記顔特徴を前記コンピュータに関連付けられたメモリに記憶するステップと、

をさらに含む、請求項 1 のプロセス。

【請求項 3】

同時捕捉されたフレームペアの新しいシーケンスを入力するステップであり、各フレームペアは、カラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む、ステップと、

プロセス動作 (b) 乃至 (e) を繰り返すステップと、

前記環境において検出され、同時捕捉されたフレームペアの前記新しいシーケンスにおいて描写された各人について、

10

人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応するかを判定するステップと、

前記人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応すると判定されるたび、前記人の同一性が事前に確定されたかを判定するステップと、

前記人の同一性が事前に確定されなかったと判定されるたび、同時捕捉されたフレームペアの前記新しいシーケンスから生成された各顔特徴について、該顔特徴が、前記人のために確立された不明人識別子に割り当てられた各顔特徴から所定程度まで異なるかを判定するステップと、

同時捕捉されたフレームペアの前記新しいシーケンスから生成された各顔特徴について、該顔特徴が、前記人のために確立された不明人識別子に割り当てられた各顔特徴から所定程度まで異なると判定されるたび、

20

前記顔特徴を前記人のために確立された不明人識別子に割り当て、前記顔特徴を前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人の同一性を確定することを試みるステップと、

前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

をさらに含む、請求項 1 のプロセス。

【請求項 4】

30

同時捕捉されたフレームペアの新しいシーケンスを入力するステップであり、各フレームペアは、カラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む、ステップと、

プロセス動作 (b) 乃至 (e) を繰り返すステップと、

前記環境において検出され、同時捕捉されたフレームペアの前記新しいシーケンスにおいて描写された各人について、

人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応するかを判定するステップと、

前記人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応すると判定されるたび、前記人の同一性が事前に確定されたかを判定するステップと、

40

前記人の同一性が事前に確定されなかったと判定されるたび、同時捕捉されたフレームペアのシーケンスが前記人の同一性を確定されることなしに何回入力され、処理されたかを決定し、その回数が所定最大数より大きいかを判定するステップと、

同時捕捉されたフレームペアのシーケンスが前記人の同一性を確定されることなしに入力され、処理された回数が前記所定最大数より大きいと判定されるたび、前記メモリから、前記人のために確立された不明人識別子に割り当てられた各顔特徴を削除するステップと、

をさらに含む、請求項 1 のプロセス。

【請求項 5】

50

同時捕捉されたフレームペアの新しいシーケンスを入力するステップであり、各フレームペアは、カラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む、ステップと、

プロセス動作 (b) 乃至 (e) を繰り返すステップと、

前記環境において検出され、同時捕捉されたフレームペアの前記新しいシーケンスにおいて描写された各人について、

人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応するかを判定するステップと、

前記人が、前記新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応しないと判定されるたび、

前記人のために生成された各顔特徴を、前記人のために確立された不明人識別子に割り当てるステップと、

前記顔特徴の各々を、前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人の同一性を確定することを試みるステップと、

前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

をさらに含む、請求項 1 のプロセス。

#### 【請求項 6】

人のために生成された各顔特徴を、前記人のために確立された不明人識別子に割り当てるプロセス動作を実行する前に、

所定最大距離より大きい前記カラービデオカメラからの距離において検出された各人について、

人の位置を、ズーム能力を有するカラーカメラを制御する制御部に提供するステップであり、前記制御部は、前記人の位置に基づいて、前記人の顔に対して、前記カラービデオカメラから前記人までの距離に比例する度合までズームインし、前記人の顔のズームされた画像を捕捉する能力がある、ステップと、

前記人の顔の前記ズームされた画像を入力するステップと、

前記人の顔を描写する前記ズームされた画像の部分の顔特徴を生成するステップと、を含むプロセス動作を実行することをさらに含む、請求項 1 のプロセス。

#### 【請求項 7】

同時捕捉されたフレームペアの追加シーケンスを入力するステップであり、各追加フレームペアは、追加のカラービデオカメラから出力されたフレームと追加の深度ビデオカメラから出力されたフレームとを含み、前記の追加のカラー及び深度ビデオカメラは、前記環境において、他のカラー及び深度ビデオカメラと同じシーンを異なる視点から捕捉し、各追加フレームペアは、前記他のカラー及び深度ビデオカメラから出力されるフレームペアと実質的に同時に捕捉される、ステップと、

顔検出方法と前記追加のカラービデオカメラからのフレームとを使用して、環境において潜在的な人を検出するステップと、

動き検出方法と前記追加の深度ビデオカメラからのフレームとを使用して、前記環境において潜在的な人を検出するステップと、

前記顔検出方法と前記動き検出方法とを介して生成された検出結果を使用して、前記環境において 1 又は複数の人の位置を決定するステップであり、前記顔検出方法を介して生成された検出結果は、検出された各人について、人の顔を描写する前記のカラービデオカメラフレームの部分の顔特徴を含む、ステップと、

単に前記動き検出方法を介して検出された各人について、

人の対応する位置を、前記追加のカラービデオカメラの同時捕捉されたフレームにおいて識別するステップと、

前記人の顔を描写する前記の追加のカラービデオカメラフレームの部分の顔特徴を生

10

20

30

40

50

成するステップと、

前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記環境において検出された各人について、

人の識別された位置に基づいて、前記人が前記他のカラー及び深度ビデオカメラを用いて同様に検出されているかどうかを判定するステップと、

前記人が前記他のカラー及び深度ビデオカメラを用いて同様に検出されていると判定されるたび、前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記人のために生成された各顔特徴を、前記他のカラー及び深度ビデオカメラを用いた前記人の検出に基づいて前記人のために確立された不明人識別子に割り当て、前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記人のために生成された前記顔特徴の各々を、前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人が前記他のカラー及び深度ビデオカメラを用いて同様に検出されていないと判定されるたび、前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記人のために生成された各顔特徴を、前記人のために確立された不明人識別子に割り当て、前記人の同一性を確定することを試み、前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

を含むプロセス動作をさらに含む、請求項 1 のプロセス。

#### 【請求項 8】

同時捕捉されたフレームペアの追加シーケンスを入力するステップであり、各追加フレームペアは、追加のカラービデオカメラから出力されたフレームと追加の深度ビデオカメラから出力されたフレームとを含み、前記追加のカラー及び深度ビデオカメラは、前記環境において他のカラー及び深度ビデオカメラと異なるシーンを捕捉する、ステップと、

顔検出方法と前記追加のカラービデオカメラからのフレームとを使用して、環境において潜在的な人を検出するステップと、

動き検出方法と前記追加の深度ビデオカメラからのフレームとを使用して、前記環境において潜在的な人を検出するステップと、

前記顔検出方法と前記動き検出方法とを介して生成された検出結果を使用して、前記環境において 1 又は複数の人の位置を決定するステップであり、前記顔検出方法を介して生成された検出結果は、検出された各人について、人の顔を描写する前記のカラービデオカメラフレームの部分の顔特徴を含む、ステップと、

単に前記動き検出方法を介して検出された各人について、

人の対応する位置を、前記追加のカラービデオカメラの同時捕捉されたフレームにおいて識別するステップと、

前記人の顔を描写する前記の追加のカラービデオカメラフレームの部分の顔特徴を生成するステップと、

前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記環境において検出された各人について、

検出された人が前記環境において異なるシーンにおいて事前に検出されたかを判定するステップと、

前記人が前記環境において異なるシーンにおいて事前に検出されていた場合、前記人の同一性が事前に確定されたかを判定するステップと、

前記人の同一性が事前に確定されなかったと判定されるたび、同時捕捉されたフレームペアの前記追加のシーケンスから生成される各顔特徴について、該顔特徴が、前記人のために事前に確立された不明人識別子に割り当てられた各顔特徴から所定程度まで異なるかを判定するステップと、

同時捕捉されたフレームペアの前記追加シーケンスから生成された各顔特徴について、該顔特徴が、前記人のために事前に確立された不明人識別子に割り当てられた顔特

10

20

30

40

50

徴から所定程度まで異なると判定されるたび、前記顔特徴を前記人のために事前に確立された不明人識別子に割り当て、前記顔特徴を前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人の同一性を確定することを試みるステップと、

前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

前記人が前記環境において異なるシーンにおいて事前に検出されなかった場合、

前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記人のために生成された各顔特徴を、前記人のために新たに確立された不明人識別子に割り当てるステップと、

前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記人のために生成された前記顔特徴の各々を、前記コンピュータに関連付けられたメモリに記憶するステップと、

前記人の同一性を確定することを試みるステップと、

前記人の同一性が確定されるたび、前記人のために確立された不明人識別子に割り当てられた各顔特徴を、前記人のために確立された顔認識トレーニングデータベースに再割り当てするステップと、

を含むプロセス動作をさらに含む、請求項 1 のプロセス。

#### 【請求項 9】

同時捕捉されたフレームペアの追加シーケンスを入力するステップであり、各追加フレームペアは、追加のカラービデオカメラから出力されたフレームと追加の深度ビデオカメラから出力されたフレームとを含み、前記追加のカラー及び深度ビデオカメラは、前記環境において他のカラー及び深度ビデオカメラと異なるシーンを捕捉する、ステップと、

顔検出方法と前記追加のカラービデオカメラからのフレームとを使用して、環境において潜在的な人を検出するステップと、

動き検出方法と前記追加の深度ビデオカメラからのフレームとを使用して、前記環境において潜在的な人を検出するステップと、

前記顔検出方法と前記動き検出方法とを介して生成された検出結果を使用して、前記環境において 1 又は複数の人の位置を決定するステップであり、前記顔検出方法を介して生成された検出結果は、検出された各人について、人の顔を描写する前記のカラービデオカメラフレームの部分の顔特徴を含む、ステップと、

単に前記動き検出方法を介して検出された各人について、

人の対応する位置を、前記追加のカラービデオカメラの同時捕捉されたフレームにおいて識別するステップと、

前記人の顔を描写する前記の追加のカラービデオカメラフレームの部分の顔特徴を生成するステップと、

前記追加のカラービデオカメラと前記追加の深度ビデオカメラとから出力されたフレームペアに基づいて前記環境において検出された各人について、

検出された人が前記環境において異なるシーンにおいて事前に検出されたかを判定するステップと、

前記人が前記環境において異なるシーンにおいて事前に検出されていた場合、

前記人の同一性が事前に確定されたかを判定するステップと、

前記人の同一性が事前に確定されなかったと判定されるたび、同時捕捉されたフレームペアの前記追加シーケンスから生成された各顔特徴について、該顔特徴が、前記人のために確立された顔認識トレーニングデータベースに割り当てられた各顔特徴から所定程度まで異なるかを判定するステップと、

同時捕捉されたフレームペアの前記追加シーケンスから生成される各顔特徴について、該顔特徴が、前記人のために確立された顔認識トレーニングデータベースに割り当てられた各顔特徴から所定程度まで異なると判定されるたび、前記顔特徴を前記人の

10

20

30

40

50

ために確立された顔認識トレーニングデータベースに割り当て、前記顔特徴を前記コンピュータに関連付けられたメモリに記憶するステップと、

を含むプロセス動作をさらに含む、請求項 1 のプロセス。

【請求項 10】

動き検出方法と前記深度ビデオカメラフレームとを使用して前記環境において潜在的な人を検出するプロセス動作は、

最初の深度ビデオカメラフレーム内のすべての画素を背景画素として指定するステップと、

同時捕捉されたフレームペアの前記シーケンスに含まれる、フレームが捕捉された順序において後から捕捉された深度フレームの各々の各画素について、

画素の深度値が、前記環境内で同じ位置を表す現在検討下にあるフレームの直前に捕捉された深度フレーム内の画素の値から所定量より大きく変化したかを識別するステップと、

前記画素の深度値が前記所定量より大きく変化したたび、前記画素を前景画素であるように指定するステップと、

一旦、同時捕捉されたフレームペアの前記シーケンスに含まれる最後のフレームが処理されて、その画素深度値が前記所定量より大きく変化したかを識別すると、

( i ) 前記最後のフレーム内の前景画素の中からシードポイントを確立し、前記シードポイントに関連付けられた画素を別個のプロブの一部であるように割り当てるステップと、

( i i ) 前記プロブに割り当てられた画素に隣接する、前記プロブにまだ割り当てられていない各画素について、その深度値が、前記プロブに割り当てられた画素の現在の平均と所定許容差内で同じであるかを再帰的に判定し、そうである場合、プロブに割り当てられておらず、かつ該プロブに割り当てられた画素の現在の平均と前記所定許容差内で同じである深度値を有する隣接する画素が見つからなくなるまで、前記隣接する画素を前記プロブの一部であるように割り当てるステップと、

( i i i ) 前記の再帰的な判定する動作 ( i i ) の実行の間、異なるプロブに割り当てられる隣接する画素が見つけれられるたび、その 2 つのプロブを 1 つに結合し、前記の再帰的な判定する動作 ( i i ) を続けるステップと、

( i v ) プロブが形成され得なくなるまで、割り当てられていない前景画素について、プロセス動作 ( i ) 乃至 ( i i i ) を繰り返すステップと、

一旦プロブが形成され得なくなると、各プロブについて、

プロブが、人間を表すプロブを示す所定基準セットに一致するかを判定するステップと、

前記所定基準セットに一致しない各プロブを消去するステップと、

各々の残りのプロブを指定して前記環境内に位置する異なる潜在的な人を表すステップと、

を含む動作を含む、請求項 1 のプロセス。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、コンピュータにより実施されるプロセスを用いて、環境において検出された各人のための顔認識トレーニングデータベースを生成することに関する。

【背景技術】

【0002】

画像内に描写された人々をその顔の見掛けから認識するという問題は、長年にわたって研究され続けている。顔認識システム及びプロセスは、基本的に、人の顔についての数種のモデルと入力画像から抽出された人の顔の画像又は特徴とを比較することによって動作する。これらの顔モデルは、通常、人の顔 ( 又はその特徴 ) の画像を用いて顔認識システムをトレーニングすることによって得られる。したがって、顔画像又は特徴をトレーニン

10

20

30

40

50

グするデータベースが、顔認識システムをトレーニングするために通常必要とされる。

【発明の概要】

【発明が解決しようとする課題】

【0003】

こうして、顔認識システムをトレーニングするために、顔画像又は特徴をトレーニングするデータベースが通常必要とされる。

【課題を解決するための手段】

【0004】

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、概して、ある時間にわたって、及び人がある環境の中を進むときに、捕捉された人の顔の特徴を収集して、その人の顔特徴のトレーニングデータベースを作成することを伴う。1つの実施形態において、コンピュータにより実施されるプロセスが採用されて、環境において検出された各人のための顔認識トレーニングデータベースを生成する。上記プロセスは、同時捕捉されたフレームペアのシーケンスを入力することで始まる。各フレームペアは、カラービデオカメラから出力されたフレームと深度ビデオカメラから出力されたフレームとを含む。次に、顔検出方法とカラービデオカメラフレームとを使用して、上記環境において潜在的な人を検出する。さらに、動き検出方法と深度ビデオカメラフレームとを使用して、上記環境において潜在的な人を検出する。

10

【0005】

前述の顔及び動き検出方法を介して生成される検出結果は、環境において1又は複数の人の位置を決定するために使用される。上記顔検出方法を介して生成される検出結果は、検出される各人について、人の顔を描写するカラービデオカメラフレームの部分の顔特徴も含む。単に動き検出方法を介して検出される各人について、上記プロセスは、その人の対応する位置をカラービデオカメラの同時捕捉されたフレームにおいて識別することと、カラービデオカメラフレームの部分の顔特徴を生成することとも含む。

20

【0006】

環境において検出される各人について、人のために生成される各顔特徴は、具体的にこの人のために確立された不明人識別子に割り当てられ、上記プロセスを実施するために使用されているコンピュータに関連付けられたメモリに記憶される。それから、各人の同一性を確定するように試みが行われる。試みがこの人に関して成功する場合、この人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる。

30

【0007】

前述の要約は、以下で詳細な説明においてさらに説明される概念のうち選択されたものを簡易な形式で案内するために提供されていることに留意されたい。この要約は、クレームされる主題の重要な特性又は必須の特性を特定するものではなく、クレームされる主題の範囲を定めることの補助として用いられるためのものでもない。

【図面の簡単な説明】

【0008】

開示の特定の特性、態様及び利点が、下記説明と添付される特許請求の範囲と関連する図面とを参照して、さらに理解されるであろう。

40

【図1A】環境において検出された各人のための顔認識トレーニングデータベースを生成する、コンピュータにより実施されるプロセスの1つの実施形態を大まかに概説するフロー図である。

【図1B】環境において検出された各人のための顔認識トレーニングデータベースを生成する、コンピュータにより実施されるプロセスの1つの実施形態を大まかに概説するフロー図である。

【図2A】環境において検出された各人のための顔認識トレーニングデータベースを、同時捕捉されたフレームペアの新しいシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの1つの実施形態を大まかに概説するフロー図である。

50



【図 2 B】環境において検出された各人のための顔認識トレーニングデータベースを、同時捕捉されたフレームペアの新しいシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 2 C】環境において検出された各人のための顔認識トレーニングデータベースを、同時捕捉されたフレームペアの新しいシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 2 D】環境において検出された各人のための顔認識トレーニングデータベースを、同時捕捉されたフレームペアの新しいシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 2 E】環境において検出された各人のための顔認識トレーニングデータベースを、同時捕捉されたフレームペアの新しいシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 3】人が、この人を識別するための所定回数の試みを超えて識別されないままであるとき、不明人識別子に割り当てられている顔特徴を破棄する、コンピュータにより実施されるプロセスの 1 つの実施形態を概説するフロー図である。

【図 4】環境において所定最大距離より大きいカラービデオカメラからの距離に位置する人のズームイン画像を捕捉する、コンピュータにより実施されるプロセスの 1 つの実施形態を概説するフロー図である。

【図 5 A】環境において検出された各人のための顔認識トレーニングデータベースを、このシーンを異なる視点から捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 5 B】環境において検出された各人のための顔認識トレーニングデータベースを、このシーンを異なる視点から捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 5 C】環境において検出された各人のための顔認識トレーニングデータベースを、このシーンを異なる視点から捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 6 A】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 6 B】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 6 C】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 6 D】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 6 E】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

10

20

30

40

50

【図 6 F】環境において検出された各人のための顔認識トレーニングデータベースを、環境内で異なるシーンを捕捉したカラー及び深度ビデオカメラの追加ペアにより出力される同時捕捉されたフレームペアのシーケンスに基づいて生成し、又は補う、コンピュータにより実施されるプロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 7 A】本明細書に説明される顔認識トレーニングデータベース生成手法実施形態における使用のための、コンピュータにより実施される動き検出プロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 7 B】本明細書に説明される顔認識トレーニングデータベース生成手法実施形態における使用のための、コンピュータにより実施される動き検出プロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 7 C】本明細書に説明される顔認識トレーニングデータベース生成手法実施形態における使用のための、コンピュータにより実施される動き検出プロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 7 D】本明細書に説明される顔認識トレーニングデータベース生成手法実施形態における使用のための、コンピュータにより実施される動き検出プロセスの 1 つの実施形態を大まかに概説するフロー図である。

【図 8】本明細書に説明される顔認識トレーニングデータベース生成手法を実施することができる適切な可動性ロボット装置の簡易コンポーネント図である。

【図 9】本明細書に説明される顔認識トレーニングデータベース生成手法を実施するための例示的システムを構成する汎用目的コンピューティング装置を描写する図である。

【発明を実施するための形態】

【0009】

下記の顔認識トレーニングデータベース生成手法実施形態の説明において、添付図面に対して参照がなされ、この添付図面は、下記の説明の一部を成し、本願手法を実施することができる特定実施形態を例示として図示する。本願手法の範囲から逸脱することなしに他の実施形態を利用することができ、構造的変更を行うことができることを理解されたい。

【0010】

さらに、明りょうさを目的として本発明の説明において特定専門用語に依ることになり、かつ本発明はそうして選ばれた特定用語に限定されることは意図されないことに留意されたい。さらに、各特定用語は、同様の目的を達成するために概して類似する方法で動作するすべてのその技術的均等物を含むことを理解されたい。本明細書において「1 つの実施形態」又は「実施形態」という表現は、実施形態と関連して説明される特定の特性、構造又は特徴が本発明の少なくとも 1 つの実施形態に含まれ得ることを意味する。本明細書の様々な箇所における「1 つの実施形態において」という表現の出現は、必ずしもすべて同じ実施形態を参照するものではなく、他の実施形態に対して相互排他的な別個又は代替の実施形態ではない。さらに、1 又は複数の実施形態を表すプロセスフローの順序は、本質的にいかなる特定順序も示すものではなく、本発明のいかなる限定も示すものではない。

【0011】

#### 1.0 顔認識システムにおける使用のためのトレーニングデータベース生成

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、概して、ある時間にわたって、及び人がある環境の中を進むときに、捕捉された人の顔の特徴を収集して、この人のための顔特徴 (facial characterizations) のトレーニングデータベースを作成することを伴う。顔特徴がある時間にわたり捕捉されるので、この顔特徴は、様々な角度及び距離、種々の解像度、並びに種々の環境条件 (例えば、照明及びヘイズ条件) 下から見られたとおりの人の顔を表すことになる。さらになお、人の顔特徴が周期的に収集される長期間の時間にわたり、これらの特徴は、人の見掛けにおける展開を表すことができる。例えば、人は、体重を増し、又は減らすことがあり、顔ひげを生やし、又は取り去ることがあり、髪形を変えることがあり、種々の帽子をかぶることがあるなどする

10

20

30

40

50

。したがって、結果として生じるトレーニングデータベースは、トレーニングが実際に始まる前に確立され、事前設定される (populated) ことができ、ある時間にわたり追加されて人の顔のポーズ及び見掛けにおける前述された変化を捕捉することができる。これは、顔認識システムにおける使用のための豊富なトレーニングリソースを生み出す。さらに、人の顔の認識トレーニングデータベースが、顔認識システムによって必要とされる前に確立されることができ、一旦採用されると、トレーニングがより迅速になることになる。さらになお、本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、環境において見つけれられた複数人のためのトレーニングデータベースを生成することができる。さらに、既存のデータベースは、顔における漸進的变化を用いて更新されることができる。これは、人の顔の特性がある時間期間にわたって大幅に変化したときでさえ、人の認識を可能にするほど十分に徐々に人の顔の変化を捕捉することを可能にする。例えば、人があごひげを生やしている場合、その顔の特性はゆっくり変化することになる。しかしながら、毎日なので、この変化は十分に小さくて、部分的なあごひげを備えた新しい顔がデータベースに追加されることができる。こうして、人のあごひげが完全に生やされたとき、手動トレーニングがこの顔を用いて行われなかったとしても、この人はなお認識されることができる。同じ原理が、年齢、体重などに起因するいかなる緩やかな変化にも当てはまる。

10

#### 【0012】

本開示を通じて使用される用語「環境」は、人の任意の外部環境として広く解釈されるべきであることに留意されたい。これは、室内設定、室外設定又は双方の組み合わせを含む。

20

#### 【0013】

##### 1.1 顔認識トレーニングデータベースを生成するプロセス

図1A～図1Bを参照し、ある環境に位置していると描写された各人のための顔認識トレーニングデータベースを生成するための、コンピュータにより実施されるプロセスの1つの実施形態が提示される。本プロセスは、同時捕捉された (contemporaneously-captured) フレームペアのシーケンスを入力することで始まる (プロセス動作100)。各フレームペアは、カラービデオカメラから出力されるフレームと深度ビデオカメラから出力されるフレームとを含む。これらカメラは、各カメラがシーンの画像を同時刻に捕捉するという点で、同期される。したがって、色フレームと深度フレームとの同時ペアが、シーンが捕捉されるたびに生み出される。次いで、顔検出方法とカラービデオカメラフレームとを使用して、環境において潜在的な人を検出する (プロセス動作102)。カラービデオフレームを採用する任意の適切な顔検出方法が、前述のタスクを遂行するために採用されてよいことに留意されたい。さらに、動き検出方法と深度ビデオカメラフレームとを使用して、環境において潜在的な人を検出する (プロセス動作104)。深度ビデオフレームを採用する任意の適切な動き検出方法が、前述のタスクを遂行するために採用されてよいことに留意されたい。1つの実装において (図1Aに例示されるとおり)、プロセス動作102と104とはほぼ同時に遂行される。

30

#### 【0014】

前述の顔及び動き検出方法を介して生成される検出結果を使用して、環境において1又は複数の人の位置を決定する (プロセス動作106)。顔検出方法を介して生成される検出結果は、検出された各々の潜在的な人について、人の顔を描写するカラービデオカメラフレームの部分の顔特徴も含む。顔特徴のタイプは、採用される特定の顔検出方法に固有であり、生成されるトレーニングデータベースを使用することになる前述された顔認識システムと互換性がある。次いで、単に動き検出方法を介して検出された各人が識別され (プロセス動作108)、各々の識別された人の対応する位置がカラービデオカメラの同時捕捉されたフレームにおいて見つけれられる (プロセス動作110)。さらに、カラービデオカメラフレームの部分の顔特徴が、識別された人の各々について生成される (プロセス動作112)。

40

#### 【0015】

50

本プロセスは、環境において検出された人のうちこれまでに選択されていない人を選択することに続く（プロセス動作 1 1 4）。この選択された人のために生成された各顔特徴が、具体的にこの人のために確立された不明人識別子（unknown person identifier）に割り当てられ（プロセス動作 1 1 6）、本プロセスを実施するために使用されているコンピュータに関連付けられたメモリに記憶される（プロセス動作 1 1 8）。前述されたコンピュータは、例えば、本開示の例示的オペレーティング環境セクションに説明されるコンピュータのうちの 1 つであってよい。

#### 【0016】

上記手順のこの時点まで、顔特徴は不明人識別子に割り当てられていることが指摘される。このように、検出された人の同一性（identity）がまだわからないとしても、顔特徴が作成され、保存されている。したがって、検出された人の同一性が最終的に確立された場合、保存された顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされることができる。この目的のため、本プロセスは、人の同一性を確定する試みに続く（プロセス動作 1 2 0）。この識別動作は、コンピュータと対話して識別する情報を提供するように不明人を招待することを含め、任意の適切な従来方法を用いて遂行される。次いで、試みが成功であったかが判定される（プロセス動作 1 2 2）。成功であった場合、選択された人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる（プロセス動作 1 2 4）。プロセス動作 1 2 0 の試みが成功であったか否かにかかわらず、次いで、すべての検出された人が選択されたかが判定される（プロセス動作 1 2 6）。選択されていない場合、すべての検出された人が選択され、検討されるまで、プロセス動作 1 1 4 乃至 1 2 6 が繰り返される。この時点で、本プロセスは終了する。

#### 【0017】

##### 1.2 後から捕捉されたフレームペアシーケンス

上記のような識別子が人に対して事前に発行されていたにもかかわらず、後から捕捉された（subsequently-captured）フレームペアのシーケンスにおいて検出された同一人物が新しい不明人識別子に関連付けられるような状況を防止するために、前述のプロセスにおいて検出された各人の位置が追跡される。任意の適切な従来追跡方法が、この目的のために採用されてよい。したがって、未来のフレームペアシーケンスの解析において、検出された人が事前に検出されていたこと、及び不明人インジケータ又は顔認識トレーニングデータベースのいずれかに関連付けられていたことが分かることになる。こうして、人のために作成された顔特徴は適切な集合に割り当てられることができ、新しい不明人識別子が確立される必要はなくなる。

#### 【0018】

前述を仮定すると、後から捕捉されたフレームペアシーケンスにおいて検出された人がどのように扱われることになるかに関して、複数の可能性が存在する。例えば、検出された人が事前に検出されており、追跡されている場合、新しいシーケンスから作成される任意の顔特徴は、この人が事前に識別されていなかった場合はこの人の既存の不明人識別子に割り当てられることになり、あるいはこの人が事前に識別されていた場合はこの人の顔認識トレーニングデータベースに割り当てられることになる。一方、検出された人がシーンに対して新しい場合、不明人識別子が作成され、生み出された顔特徴に割り当てられることになる。追加的に、顔特徴が不明人インジケータ（既存のもの又は新しいものにかかわらず）に割り当てられるたび、この人を識別する試みが行われることになる。

#### 【0019】

より詳細に、図 2 A ~ 図 2 E を参照すると、1 つの実施形態において、同時捕捉されたフレームペアの新しいシーケンスが利用可能になるとき、該シーケンスが入力される（プロセス動作 2 0 0）。それから、図 1 のプロセス動作 1 0 2 乃至 1 1 2 が、新しいフレームペアシーケンスを用いて行われる。

#### 【0020】

それから、本プロセスは、新しいフレームペアシーケンスを用いて、環境において検出

された人のうち1人を選択することに続く(プロセス動作202)。それから、選択された人が、新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応するかが判定される(プロセス動作204)。これまでに示されたとおり、1つの実施形態において、これは、事前に検出された人の位置をある時間にわたり追跡することによって行われる。人が、こうした事前に検出された人に対応すると判定される場合、次いで、この人の同一性が事前に確定されたかが判定される(プロセス動作206)。この人の同一性が事前に確定されていた場合、この人のための同時捕捉されたフレームペアの新しいシーケンスから生成された顔特徴のうち事前に選択されていない1つが選択される(プロセス動作208)。顔特徴は、これまでに説明されたように生成されることに留意されたい。選択された顔特徴が、この人のために確立された顔認識トレーニングデータベースに割り当てられた各顔特徴と所定程度まで異なるかが判定される(プロセス動作210)。選択された顔特徴が所定程度まで異なる場合、この選択された顔特徴は、選択された人のために確立された顔認識トレーニングデータベースに割り当てられ(プロセス動作212)、コンピュータに関連付けられたメモリに記憶される(プロセス動作214)。そうでなければ、この選択された顔特徴は破棄される(プロセス動作216)。いずれにしても、それから、選択された人のために新しいフレームペアシーケンスから作成されたすべての顔特徴が選択されたかが判定される(プロセス動作218)。選択されていない場合、すべての顔特徴が選択され、検討されるまで、プロセス動作208乃至218が繰り返される。

10

20

30

40

50

#### 【0021】

しかしながら、選択された人の同一性が事前に確定されていなかったとプロセス動作206において判定される場合、この人のために同時捕捉されたフレームペアの新しいシーケンスから生成された顔特徴のうち事前に選択されていない1つが選択される(プロセス動作220)。それから、選択された顔特徴が、この人のために確立された不明人識別子に割り当てられた各顔特徴と所定程度まで異なるかが判定される(プロセス動作222)。選択された顔特徴が所定程度まで異なる場合、この選択された顔特徴は選択された人のために確立された不明人識別子に割り当てられ(プロセス動作224)、コンピュータに関連付けられたメモリに記憶される(プロセス動作226)。そうでなければ、この選択された顔特徴は破棄される(プロセス動作228)。いずれにしても、それから、選択された人のために新しいフレームペアシーケンスから作成されたすべての顔特徴が選択されたかが判定される(プロセス動作230)。選択されていない場合、すべての顔特徴が選択され、検討されるまで、プロセス動作220乃至230が繰り返される。それから、本プロセスは、人の同一性を確定する試みに続く(プロセス動作232)。以前のとおりに、この識別動作は、コンピュータと対話して識別する情報を提供するように不明人を招待することを含め、任意の適切な従来方法を用いて遂行される。次いで、試みが成功であったかが判定される(プロセス動作234)。成功であった場合、選択された人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる(プロセス動作236)。

#### 【0022】

選択された人が環境に対して新しく、あるいは過去に検出されなかった可能性もある。この目的のため、選択された人が、上記の新しいシーケンスの前の同時捕捉されたフレームペアのシーケンスを用いて位置が事前に決定された人に対応しないとプロセス動作204において判定される場合、選択された人のために生成された各顔特徴は、具体的にこの人のために確立される不明人識別子に割り当てられ(プロセス動作238)、本プロセスを実施するために使用されているコンピュータに関連付けられたメモリに記憶される(プロセス動作240)。次いで、この人の同一性を確定するように試みが行われる(プロセス動作242)。それから、試みが成功であったかが判定される(プロセス動作244)。成功であった場合、選択された人のために確立された不明人識別子に割り当てられた各顔特徴が、この人のために確立される顔認識トレーニングデータベースに再割り当てされる(プロセス動作246)。

## 【 0 0 2 3 】

一旦現在選択されている人が上記で概説されたとおり検討されると、新しいフレームペアシーケンスを用いて環境において検出されたすべての人が選択されたかが判定される（プロセス動作 2 4 8）。選択されていない場合、すべての検出された人が選択され、検討されるまで、プロセス動作 2 0 2 乃至 2 4 8 が繰り返される。この時点で、プロセスの現在の反復が終了する。しかしながら、本プロセスは、同時捕捉されたフレームペアの新しいシーケンスが次に利用可能になるとき、繰り返されることができる。

## 【 0 0 2 4 】

顔認識方法は、通常、人をその顔の画像から識別することにおいて、これまでに説明された顔特徴などの顔特徴を使用する。人の同一性を確定することを試みるための前述のプロセス動作に関して、この人のために生成され、この人の不明人識別子に割り当てられた顔特徴を、上記試みにおいて採用できることに留意されたい。

## 【 0 0 2 5 】

1 . 2 . 1 同一性を確認できない人

前述のプロセスにおいて、検出された人がまったく識別されないことになる可能性もある。メモリ空間を節約するために、図 3 に概説されるような 1 つの実施形態において、プロセス動作 1 2 2、2 3 4 又は 2 4 4 のいずれか 1 つにおいて選択された人の同一性が確定されなかったと判定される場合、同時捕捉されたフレームペアのシーケンスが、この人の同一性が確定されることなしに入力され、処理された回数が、勘定される（プロセス動作 3 0 0）。それから、勘定された数が所定最大数、例えば 1 0 0 より大きいかが判定される（プロセス動作 3 0 2）。大きくない場合、上記で概説されたプロセスが停止されずに続き、このメモリ節約手順は終了する。しかしながら、勘定された数が所定最大数より大きい場合、選択された人のために確立された不明人識別子に割り当てられた各顔特徴は、コンピュータのメモリから削除される（プロセス動作 3 0 4）。

## 【 0 0 2 6 】

1 . 2 . 2 ズームスキーム

本明細書に説明される顔認識トレーニングデータベース生成手法と関連して採用できる多くの顔認識方法は、環境において検出されるがカメラから著しい距離にいる人をしばしば認識できないであろうことに留意されたい。必ずしもそうではないが、前述の状況は、人が単に動き検出方法を介して検出されるとき、発生することがある。この状況には、ズームスキームを用いて対処することができる。このズームスキームは、人のために生成された各顔特徴を、この人のために確立された不明人識別子に割り当てる前に、遂行される。より詳細には、図 4 を参照すると、1 つの実施形態において、環境に存在していると（前述された方法のうちいずれかによって）検出された、事前に選択されていない人が、選択される（プロセス動作 4 0 0）。それから、選択された人が、環境において、所定最大距離、例えば 3 メートルより大きいカラービデオカメラからの距離にいるかが判定される（プロセス動作 4 0 2）。大きい場合、選択された人の位置が、ズーム能力を有するカラーカメラを制御する制御部に提供される（プロセス動作 4 0 4）。制御部はカラーカメラに、カラービデオカメラからこの人までの距離に比例する度合で、選択された人の顔に対してズームインさせ、それから人の顔のズームされた画像を捕捉させる。カラーカメラは、環境の画像を捕捉するように位置付けられた前述のカラービデオカメラ又は別個のカメラであってよいことに留意されたい。ズームの度合は、カメラから選択された人までの距離を所与として、結果的に生じる画像が顔認識の助けとなる解像度で人の顔を描写することになるように、計算される。それから、ズームされた画像は入力され（プロセス動作 4 0 6）、この人の顔を描写するズームされた画像の部分の顔特徴が生成される（プロセス動作 4 0 8）。それから、この顔特徴が、選択された人のために生成されたすべての他のものと共に、この人のために確立された不明人識別子に割り当てられる。

## 【 0 0 2 7 】

1 . 3 追加のカラー及び深度ビデオカメラ

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態が動作する環

10

20

30

40

50

境は、かなり大きくなり得る。そのようなものとして、1つの実施形態において、2ペア以上のカラー及び深度ビデオカメラが、環境をカバーするために採用される。2ペア以上のカメラが環境において利用可能であると仮定して、これらカメラは、同じシーンを種々の視点から捕捉するように構成されることができる。このシナリオは、種々のペアのカメラによって検出される同じ人々について、又は、あるペアのカメラが「見る」ことができない人を別のカメラが「見る」ことができるときに異なる人々について、より多くの顔特徴が同じ時間期間内に生成されることを可能にする。これに関し、各ペアのカメラがシーンの中の人々の位置を知ることは有利であり、その結果、ある人が別のカメラペアを用いて検出された同一人物であるか、あるいは異なる人であるかを容易に判定することができる。1つの実施形態において、これは、フレームペアを実質的に同時に捕捉するようにカメラペアを構成することによって遂行される。こうして、カメラの1ペアにより計算される人の位置は、この人が同一人物である場合、別のペアにより計算される位置に一致することになり、この人が異なる人である場合、一致しないことになる。

10

20

30

40

50

#### 【0028】

##### 1.3.1 同じシーンを異なる視点から捕捉する

より詳細に、図5A～図5Cを参照すると、異なる視点からシーンを捕捉するカラー及び深度ビデオカメラの各追加ペアについて、同時捕捉されたフレームペアの追加シーケンスが入力される（プロセス動作500）。次いで、顔検出方法とカメラの追加ペアのうちカラービデオカメラにより出力されたカラービデオカメラフレームとを使用して、環境において潜在的な人を検出する（プロセス動作502）。さらに、動き検出方法とカメラの追加ペアのうち深度ビデオカメラにより出力された深度ビデオカメラフレームとを使用して、環境において潜在的な人を検出する（プロセス動作504）。前述の顔及び動き検出方法を介して生成された検出結果を使用して、環境において1又は複数の人の位置を決定する（プロセス動作506）。顔検出方法を介して生成された検出結果は、検出された各々の潜在的な人について、人の顔を描写するカラービデオカメラフレームの部分の顔特徴も含む。

#### 【0029】

次いで、単に動き検出方法を介して検出された各人が識別され（プロセス動作508）、各々の識別された人の対応する位置が、カメラの追加ペアのうちカラービデオカメラの同時捕捉されたフレームにおいて見つけられる（プロセス動作510）。さらに、カラービデオカメラフレームの上記部分の顔特徴が、識別された人の各々について生成される（プロセス動作512）。

#### 【0030】

本プロセスは、追加のカラー及び深度ビデオカメラペアから出力されたフレームペアに基づいて環境において検出された人のうち事前に選択されていない1人を選択することに基づく（プロセス動作514）。それから、この人の識別された位置に基づいて、この人が別のカラー及び深度ビデオカメラペアを用いて同様に検出されているかどうか判定される（プロセス動作516）。検出されている場合、追加のカラー及び深度ビデオカメラペアから出力されたフレームペアに基づいてこの選択された人のために生成された各顔特徴は、他のカラー及び深度ビデオカメラペアを用いたこの人の検出に基づいてこの人のために確立された不明人識別子に割り当てられる（プロセス動作518）。そうでなければ、追加のカラー及び深度ビデオカメラペアから出力されたフレームペアに基づいてこの選択された人のために生成された各顔特徴は、この人のために確立される不明人識別子に割り当てられる（プロセス動作520）。いずれにせよ、追加のカラー及び深度ビデオカメラペアから出力されたフレームペアに基づいてこの選択された人のために生成された顔特徴の各々は、コンピュータに関連付けられたメモリに記憶される（プロセス動作522）。さらに、この人の同一性を確定するように試みも行われる（プロセス動作524）。それから、この試みが成功であったかが判定される（プロセス動作526）。成功であった場合、この選択された人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる（プロセス

動作 5 2 8 )。プロセス動作 5 2 6 の試みが成功であったか否かにかかわらず、次いで、すべての検出された人が選択されたかが判定される (プロセス動作 5 3 0)。選択されていない場合、すべての検出された人が選択され、検討されるまで、プロセス動作 5 1 4 乃至 5 3 0 が繰り返される。この時点で本プロセスは終了するが、同時捕捉されたフレームペアの新しいシーケンスがカラー及び深度ビデオカメラの追加ペアから入力されるたび、繰り返されることができる。

#### 【 0 0 3 1 】

##### 1 . 3 . 2 異なるシーンを捕捉する

2 ペア以上のカメラが環境において利用可能であると仮定して、このカメラペアは、種々のシーンを捕捉するように構成することができる場合もある。この構成は、カメラのペアが環境全体をカバーできない状況において有用である。これを仮定すると、あるカメラペアがカバーするあるシーンにおいて検出された人を追跡することができ、この人が、別のカメラペアがカバーする環境の部分へ動く場合、この人があるシーンから別のシーンへと去ったときのこの人の位置についての知識を使用して、新しいシーンにおいて検出される人が前のシーンにおいて検出された同一人物であることを確定することができる。さらに、新しいシーンにおいて検出された人が前のシーンにおいて検出された同一人物であると確定することが実現可能な場合、顔認識方法、又はこの人を識別する何らかの他の方法が採用されてよい。これは、環境の新しい部分においてこの人のために生成された顔特徴を、正しい不明人識別子 (又は、この人が事前に識別されている場合は正しい顔認識トレーニングデータベース) に割り当てることを容易にする。

10

20

#### 【 0 0 3 2 】

より詳細に、図 6 A ~ 図 6 F を参照すると、環境内で異なるシーンを捕捉するカラー及び深度ビデオカメラの追加ペアを所与として、同時捕捉されたフレームペアの追加シーケンスが入力される (プロセス動作 6 0 0)。次いで、顔検出方法とカメラの追加ペアのうちカラービデオカメラにより出力されたカラービデオカメラフレームとを使用して、環境において潜在的な人を検出する (プロセス動作 6 0 2)。さらに、動き検出方法とカメラの追加ペアのうち深度ビデオカメラにより出力された深度ビデオカメラフレームとを使用して、環境において潜在的な人を検出する (プロセス動作 6 0 4)。前述の顔及び動き検出方法を介して生成された検出結果を使用して、環境において 1 又は複数の人の位置を決定する (プロセス動作 6 0 6)。顔検出方法を介して生成された検出結果は、検出された各々の潜在的な人について、人の顔を描写するカラービデオカメラフレームの部分の顔特徴も含む。

30

#### 【 0 0 3 3 】

次いで、単に動き検出方法を介して検出された各人が識別され (プロセス動作 6 0 8)、各々の識別された人の対応する位置が、カメラの追加ペアのうちカラービデオカメラの同時捕捉されたフレームにおいて見つけられる (プロセス動作 6 1 0)。さらに、カラービデオカメラフレームの上記部分の顔特徴が、識別された人の各々について生成される (プロセス動作 6 1 2)。

#### 【 0 0 3 4 】

本プロセスは、追加のカラー及び深度ビデオカメラペアから出力されたフレームペアに基づいて環境において検出された人のうち事前に選択されていない 1 人を選択することに基づく (プロセス動作 6 1 4)。それから、この選択された人が別のカラー及び深度ビデオカメラペアを用いて環境において別のシーンで事前に検出されていたかどうか判定される (プロセス動作 6 1 6)。これまでに示されたとおり、これは、この人があるシーンから別のシーンへ去るときに人の位置の追跡すること、顔認識方法、又はこの人を識別する何らかの他の方法に基づくことができる。選択された人が別のシーンにおいて事前に検出されていた場合、この選択された人の同一性が事前に確定されたかがさらに判定される (プロセス動作 6 1 8)。選択された人が事前に識別されていなかった場合、同時捕捉されたフレームペアの追加シーケンスから生成された顔特徴のうち事前に選択されていない 1 つが選択され (プロセス動作 6 2 0)、選択された顔特徴が、この選択された人のために

40

50



事前に確立された不明人識別子に割り当てられた各顔特徴と所定程度まで異なるかが判定される（プロセス動作 6 2 2）。そうである場合、上記選択された顔特徴は、この人のために事前に確立された不明人識別子に割り当てられ（プロセス動作 6 2 4）、コンピュータに関連付けられたメモリに記憶される（プロセス動作 6 2 6）。そうでなければ、上記選択された顔特徴は破棄される（プロセス動作 6 2 8）。それから、同時捕捉されたフレームペアの追加シーケンスから生成されたすべての顔特徴が選択されたかが判定される（プロセス動作 6 3 0）。選択されていない場合、すべての顔特徴が選択され、検討されるまで、プロセス動作 6 2 0 乃至 6 3 0 が繰り返される。次いで、この選択された人の同一性を確定するように試みが行われる（プロセス動作 6 3 2）。それから、この試みが成功であったかが判定される（プロセス動作 6 3 4）。成功であった場合、この選択された人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる（プロセス動作 6 3 6）。

10

#### 【 0 0 3 5 】

しかしながら、プロセス動作 6 1 8 において、選択された人が事前に識別されたと判定される場合、同時捕捉されたフレームペアの追加シーケンスから生成される顔特徴のうち事前に選択されていない 1 つが選択され（プロセス動作 6 3 8）、選択された顔特徴が、選択された人のために事前に確立された顔認識トレーニングデータベースに割り当てられた各顔特徴と所定程度まで異なるかが判定される（プロセス動作 6 4 0）。そうである場合、選択された顔特徴は、この人のために確立された顔認識トレーニングデータベースに割り当てられ（プロセス動作 6 4 2）、コンピュータに関連付けられたメモリに記憶される（プロセス動作 6 4 4）。そうでなければ、選択された顔特徴は破棄される（プロセス動作 6 4 6）。それから、同時捕捉されたフレームペアの追加シーケンスから生成されるすべての顔特徴が選択されたかが判定される（プロセス動作 6 4 8）。選択されていない場合、すべての顔特徴が選択され、検討されるまで、プロセス動作 6 3 8 乃至 6 4 8 が繰り返される。

20

#### 【 0 0 3 6 】

しかしながら、プロセス動作 6 1 6 において、選択された人が環境において別のシーンで事前に検出されなかったと判定された場合、本プロセスは、追加のカラービデオカメラと追加の深度ビデオカメラとから出力されたフレームペアに基づいてこの選択された人のために生成された各顔特徴を、この人のために新たに確立される不明人識別子に割り当てることに続く（プロセス動作 6 5 0）。これらの顔特徴の各々もまた、コンピュータに関連付けられたメモリに記憶される（プロセス動作 6 5 2）。次いで、この選択された人の同一性を確定する試みが行われる（プロセス動作 6 5 4）。それから、この試みが成功であったかが判定される（プロセス動作 6 5 6）。選択された人の同一性が確定された場合、この人のために確立された不明人識別子に割り当てられた各顔特徴は、この人のために確立された顔認識トレーニングデータベースに再割り当てされる（プロセス動作 6 5 8）。

30

#### 【 0 0 3 7 】

一旦選択された人が上記で説明されたとおり検討されると、すべての検出された人が選択されたかが判定される（プロセス動作 6 6 0）。選択されていない場合、すべての検出された人が選択され、検討されるまで、プロセス動作 6 1 4 乃至 6 6 0 が繰り返される。この時点で本プロセスは終了するが、同時捕捉されたフレームペアの新しいシーケンスがカラー及び深度ビデオカメラの追加ペアから入力されるたび、繰り返されることができる。

40

#### 【 0 0 3 8 】

##### 1 . 4 動き検出

任意の動き検出方法が、本明細書に説明される顔認識トレーニングデータベース生成手法実施形態における使用のために採用されてよいが、1つの実施形態において、下記方法が採用される。概して、この方法は、深度ビデオカメラフレームから抽出される深度データにおける短期的変化を生かして、環境において潜在的な人を検出する。

50

## 【 0 0 3 9 】

より詳細に、図 7 A ~ 図 7 D を参照すると、1 つの実施形態において、動き検出プロセスは、まず、最初の深度ビデオカメラフレーム内のすべての画素を背景画素として指定することを伴う（プロセス動作 7 0 0）。それから、後から捕捉された新しい深度フレームが利用可能になったかが判定される（プロセス動作 7 0 2）。なっていない場合、新しいフレームが利用可能になるまで、プロセス動作 7 0 2 が繰り返される。新しい深度フレームが入力されると、この深度フレームのうち事前に選択されていない画素が選択され（プロセス動作 7 0 4）、選択された画素の深度値が、環境内で同一位置を表す現在検討下にあるフレームの直前に捕捉された深度フレーム内の画素の値から所定量を超えて変わったかが判定される（プロセス動作 7 0 6）。深度値が所定量を超えて変わった場合、選択された画素は、前景画素であるように指定される（プロセス動作 7 0 8）。次いで、残りの深度フレームのうちいくらかの事前に選択されていない画素があるかが判定される（プロセス動作 7 1 0）。残りの画素がある場合、プロセス動作 7 0 4 乃至 7 1 0 が繰り返される。ない場合、それから、現在検討下にある深度フレームがシーケンスにおける最後のフレームであるかが判定される（プロセス動作 7 1 2）。そうでない場合、プロセス動作 7 0 2 乃至 7 1 2 が繰り返される。

## 【 0 0 4 0 】

しかしながら、現在検討下にある深度フレームが最後のフレームである場合、シードポイント（seed point）が最後のフレーム内の前景画素の中から確立され、このポイントに関連付けられる画素はプロブ（blob）の部分であるように割り当てられる（プロセス動作 7 1 4）。次いで、プロブ（最初、単にシードポイント画素であろう）に割り当てられた画素に隣接し、かつ上記プロブにまだ割り当てられていない、事前に選択されていない画素が選択される（プロセス動作 7 1 6）。まず、選択された画素が異なるプロブに割り当てられるかが判定される（プロセス動作 7 1 8）。そうである場合、この 2 つのプロブは 1 つのプロブに結合される（プロセス動作 7 2 0）。次いで、結合されたプロブに割り当てられた画素に隣接し、かつ結合されたプロブにまだ割り当てられていない、いくらかの事前に選択されていない画素があるかが判定される（プロセス動作 7 2 2）。ある場合、これらの画素のうち事前に選択されていない 1 つが選択され（プロセス動作 7 2 4）、プロセス動作 7 1 8 乃至 7 2 4 が繰り返される。しかしながら、選択された画素が異なるプロブに割り当てられなかったとプロセス動作 7 1 8 において判定されるたび、選択された画素の深度値が、プロブに割り当てられた画素の現在の平均と所定許容差内で同じであるかが判定される（プロセス動作 7 2 6）。そうである場合、選択された画素はプロブに割り当てられる（プロセス動作 7 2 8）。そうでない場合、何の動作もとられない。しかしながら、いずれにせよ、次いで、プロブ（結合されており、又は結合されていない）に割り当てられた画素に隣接し、かつこのプロブにまだ割り当てられていない、いくらかの事前に選択されていない画素があるかが判定される（プロセス動作 7 3 0）。こうした画素がある場合、プロセス動作 7 1 6 乃至 7 3 0 が繰り返される。そうでなければ、何の動作もとられない。こうして、シードポイント画素の周りの画素は、各々検討され、プロブが結合されること、又は画素が必要な深度値を有する場合にプロブに割り当てられることをもたらし、そして、（結合され、あるいはその他の方法で）拡大されるプロブの周りの画素が検討されるなどしてプロブを大きくする。これは、プロブに割り当てられておらず、かつこのプロブに割り当てられた画素の現在の平均と上記所定許容差内で同じである深度値を有する隣接画素が見つからなくなるまで続く。

## 【 0 0 4 1 】

次いで、プロブに割り当てられていない前景画素があるかが判定される（プロセス動作 7 3 2）。こうした画素が残っている場合、シードポイントが最後のフレーム内の割り当てられていない前景画素の中から確立され、このポイントに関連付けられる画素は、新しいプロブの部分であるように割り当てられる（プロセス動作 7 3 4）。それから、割り当てられていない前景画素が残らなくなるまで、プロセス動作 7 1 6 乃至 7 3 4 が繰り返される。

10

20

30

40

50

## 【 0 0 4 2 】

一旦割り当てられていない前景画素が残らずなくなる（ゆえに、新しいプロブを形成することができなくなる）と、プロブのうち事前に選択されていない1つが選択される（プロセス動作736）。それから、このプロブが、人間を表すプロブを示す所定基準セットに一致するかが判定される（プロセス動作738）。一致しない場合、このプロブは消去される（プロセス動作740）。しかしながら、選択されたプロブが所定基準に一致する場合、このプロブは、環境内に位置する潜在的な人を表すとして指定される（プロセス動作742）。

## 【 0 0 4 3 】

人間を表すプロブを示すために使用される基準は、任意の従来の基準セットであってよいことに留意されたい。さらに、この基準は、プロブが、実空間次元における通常の人体パラメータに適合するかどうかを含むことができる。例えば、プロブは、人間の胸部及び頭部に対応する矩形領域を示す。

## 【 0 0 4 4 】

## 2.0 カラー及び深度ビデオカメラ

次に、本明細書に説明される顔認識トレーニングデータベース生成手法実施形態により採用される前述されたカラービデオカメラ及び深度ビデオカメラを、より詳細に説明していく。概して、カラービデオカメラは、該カメラにより捕捉されたシーンのデジタルカラー画像の連続的シーケンスを出力する。これらの画像は、これまでの説明にあったとおり、フレーム又は画像フレームと呼ばれることがある。適切なカラービデオカメラの一例は、従来のRGBビデオカメラである。深度ビデオカメラは、該カメラにより捕捉されたシーンのデジタル深度画像の連続的シーケンスを出力する。これらの画像は、これまでの説明にあったとおり、フレーム又は深度フレームと本明細書において呼ばれることがある。深度フレーム内の画素値は、深度カメラと環境内の被写体との間の距離を示す。例えば、1つの適切な深度ビデオカメラは、従来の赤外線ベースの深度カメラである。このタイプのカメラは、既知の赤外線パターンを環境に投影し、赤外線画像装置により捕捉されたときのパターンの変形に基づいて深度を決定する。

## 【 0 0 4 5 】

これまでに説明されたとおり、本明細書に説明される顔認識トレーニングデータベース生成手法の実施形態は、色フレームと深度フレームとの同時捕捉されたペア間の画素相関を使用することができる。換言すると、ペアのフレームのうち一方におけるどの画素が他方のフレームにおける所与の画素とシーン内で同じ位置を描写するかを知ることが、時に有用である。従来方法を採用して、同時フレームのペアが捕捉されるたびにこの画素相関を確定することができるが、1つの実施形態において、画素相関を定義する事前計算された変換が採用される。より詳細には、カラービデオカメラと深度ビデオカメラとが、これらが同じ方法で一緒に動かされるように同期される場合、これらの間の相対的変換は変わらないことになる。そのようなものとして、この変換を事前計算し、使用して、捕捉された同時フレームの各ペアのための画素相関を決定することができる。

## 【 0 0 4 6 】

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、固定された位置のカラー及び深度ビデオカメラを採用することもできる。固定された位置によって、これらカメラは、環境内で特定位置に配置され、その位置から独自に動かないことが意図される。これは、当然ながら、これらカメラを環境内で移動することを除外するものではない。しかしながら、これらカメラは、動作の間、同じ位置にとどまることが想定される。さらに、固定された位置のカメラは位置を動かないが、これは、これらカメラがその位置にいる間、パンし、チルトし、回転し、あるいはズームすることができないことを意味するものではない。

## 【 0 0 4 7 】

別法として、本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、動くカラー及び深度ビデオカメラを採用することができる。例えば、これらカメラは

、可動性ロボット装置に取り付けられることができる。適切な可動性ロボット装置は、概して、下記属性を示す任意の従来の可動性ロボット装置であってよい。まず、図 8 を参照すると、ロボット装置 800 は、該装置が進むことが意図される環境を動き回ることができる。したがって、可動性ロボット装置 800 は、環境の至る所に当該装置を動かすための運動セクション 802 を含む。可動性ロボット装置 800 は、適用できる環境の至る所で人々を追跡し観察するために採用されるセンサも有する。具体的に、これらのセンサには、前述されたカラービデオカメラ 804 と深度ビデオカメラ 806 とを含む。カラー及び深度ビデオカメラ 804、806 は位置変更可能であり、したがって、環境の種々の部分を捕捉することができる。この目的のため、カラー及び深度ビデオカメラ 804、806 は、可動性ロボット装置 800 のうち、前述された運動セクション 802 の上部に通常配置される頭部セクション 808 に収容されてよい。カメラ 804、806 の視点は、カメラ自体を向け直すことによって若しくは頭部セクション 808 を動かすことによって、又は双方で、変更することができる。後半のシナリオの一例は、カメラが上下に回転して (pivot) チルトする動きを提供しながら、頭部セクションが垂直軸に関して回転して 360 度のパンする動きを提供する構成である。これらカメラはズーム機能も有する。

10

20

30

40

50

#### 【0048】

可動性ロボット装置 800 は制御ユニット 810 も含み、制御ユニット 810 は、従来の方法で環境の至る所にロボット装置を動かすように運動セクション 802 を制御し、環境内で種々のシーンを捕捉するように頭部セクション 808 若しくはカメラ 804、806 又は双方の動きを制御する。さらに、制御ユニット 810 は、コンピューティング装置 812 (本開示の例示的オペレーティング環境セクションに説明されるコンピューティング装置など)を含む。このコンピューティング装置 812 は制御モジュールを含み、この制御モジュールは、運動セクション及び頭部セクションに対して動き制御信号を伝えることについて、及び顔認識トレーニングデータベースを生成するためにこれまでに説明された方法でカラー及び深度ビデオカメラにより捕捉されたフレームを用いることについて責任を負う。運動セクション及び頭部セクションの動きの制御は、従来方法を用いて行われる。これに対し、後半の機能は、顔認識トレーニング生成サブモジュールによって扱われる。

#### 【0049】

動作において、図 7A ~ 図 7D と関連してこれまでに説明された動き検出プロセスは、可動性ロボット装置が固定されており、かつカメラが動かない (例えば、パン、チルト、回転又はズームをしない) ときに行われるであろうことに留意されたい。これは、カメラの相対的動きに起因した誤検出を防止する。

#### 【0050】

##### 3.0 例示的オペレーティング環境

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、多数のタイプの汎用目的又は専用目的のコンピューティングシステム環境又は構成内で動作可能である。図 9 は、汎用目的コンピュータシステムの簡素化された例を示し、このコンピュータシステム上で、顔認識トレーニングデータベース生成手法の様々な実施形態と要素とを、本明細書に説明されたとおり、実施することができる。図 9 において破線又は点線により表される任意のボックスは簡素化されたコンピューティング装置の代替の実施形態を表すこと、及び、これらの代替の実施形態のうち任意のもの又はすべてを下記で説明されたとおり本明細書にわたって説明された他の代替の実施形態と組み合わせて使用できることに留意されたい。

#### 【0051】

例えば、図 9 は、簡素化されたコンピューティング装置 10 を示す汎用システム図を示す。こうしたコンピューティング装置は、通常、少なくともいくつかの最小計算能力を有する装置において見られることがあり、該装置は、これらに限定されないが、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルドコンピューティング装置、ラップトップ又はモバイルコンピュータ、携帯電話及び PDA などの通信装置、マルチプロセッサシ

ステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラム可能家電製品、ネットワークＰＣ、ミニコンピュータ、メインフレームコンピュータ、オーディオ又はビデオメディアプレーヤなどを含む。

#### 【 0 0 5 2 】

本明細書に説明される顔認識トレーニングデータベース生成手法実施形態を装置が実施することを可能にするために、該装置は、基本的な計算オペレーションを可能にするほど十分な計算能力とシステムメモリとを有するべきである。具体的には、図 9 に例示されるとおり、計算能力は、1 又は複数のプロセッシングユニット 1 2 により概して示され、1 又は複数の G P U 1 4 も含むことができ、いずれか又は双方がシステムメモリ 1 6 と通信することができる。汎用コンピューティング装置のプロセッシングユニット 1 2 は、D S P、V L I W 又は他のマイクロコントローラなどの特化されたマイクロプロセッサであってよく、あるいはマルチコア C P U の中に特化された G P U ベースのコアを含む、1 又は複数のプロセッシングコアを有する従来の C P U であってよいことに留意されたい。

10

#### 【 0 0 5 3 】

さらに、図 9 の簡素化されたコンピューティング装置は、例えば通信インタフェース 1 8 などの他のコンポーネントを含んでもよい。図 9 の簡素化されたコンピューティング装置は、1 又は複数の従来のコンピュータ入力装置 2 0 (例えば、ポインティング装置、キーボード、オーディオ入力装置、ビデオ入力装置、触覚入力装置、有線又は無線のデータ送信信号を受信する装置など)を含むこともできる。図 9 の簡素化されたコンピューティング装置は、例えば、1 又は複数の従来の表示装置 2 4 と他のコンピュータ出力装置 2 2 (例えば、オーディオ出力装置、ビデオ出力装置、有線又は無線のデータ送信信号を送信する装置など)となどの他の任意的コンポーネントを含むこともできる。汎用目的コンピュータのための典型的な通信インタフェース 1 8、入力装置 2 0、出力装置 2 2 及び記憶装置 2 6 は、当業者に周知であり、本明細書において詳細に説明されないことに留意されたい。

20

#### 【 0 0 5 4 】

図 9 の簡素化されたコンピューティング装置は、様々なコンピュータ読取可能媒体も含むことができる。コンピュータ読取可能媒体は、コンピュータ 1 0 により記憶装置 2 6 を介してアクセスすることができる任意の利用可能な媒体であってよく、コンピュータ読取可能又はコンピュータ実行可能な命令、データ構造、プログラムモジュール又は他のデータなどの情報を記憶するための取外し可能 2 8 及び / 又は取外し不能 3 0 である揮発性媒体と不揮発性媒体との双方を含む。限定ではなく例として、コンピュータ読取可能媒体は、コンピュータ記憶媒体と通信媒体とを含んでよい。コンピュータ記憶媒体は、これらに限定されないが、D V D、C D、フロッピー (登録商標) ディスク、テープドライブ、ハードドライブ、光学ドライブ、ソリッドステートメモリ装置、R A M、R O M、E E P R O M、フラッシュメモリ若しくは他のメモリテクノロジー、磁気カセット、磁気テープ、磁気ディスク記憶装置若しくは他の磁気記憶装置、又は所望される情報を記憶するために使用することができ、かつ 1 若しくは複数のコンピューティング装置によりアクセスすることができる任意の他の装置などの、コンピュータ読取可能又はマシン読取可能な媒体又は記憶装置を含む。

30

40

#### 【 0 0 5 5 】

コンピュータ読取可能又はコンピュータ実行可能な命令、データ構造、プログラムモジュールなどの情報の保有は、1 若しくは複数の変調されたデータ信号若しくは搬送波をエンコードする様々な前述された通信媒体、又は他の送信メカニズム若しくは通信プロトコルのうち、任意のものをを用いて遂行されてよく、任意の有線の又は無線の情報配信メカニズムを含む。用語「変調されたデータ信号」又は「搬送波」は、概して、その特徴のうち 1 又は複数、信号内に情報をエンコードする方法で設定され又は変更された信号を指すことに留意されたい。例えば、通信媒体は、1 又は複数の変調されたデータ信号を担体する有線ネットワーク又は直接有線接続などの有線媒体と、音波、R F、赤外線、及び 1 又は複数の変調されたデータ信号又は搬送波を送信及び / 又は受信するための他の無線媒体

50

などの無線媒体とを含む。上記のうち任意のものの組み合わせもまた、通信媒体の範囲内に含まれるべきである。

【 0 0 5 6 】

さらに、本明細書に説明される様々な顔認識トレーニングデータベース生成手法実施形態のうちいくつか又はすべてを具現化するソフトウェア、プログラム及び／又はコンピュータプログラム製品、あるいはこれらの部分は、コンピュータ実行可能命令又は他のデータ構造の形式で、コンピュータ読取可能又はマシン読取可能な媒体又は記憶装置と通信媒体との任意の所望される組み合わせから記憶され、受信され、送信され、あるいは読み出されることができる。

【 0 0 5 7 】

最後に、本明細書に説明される顔認識トレーニングデータベース生成手法実施形態は、コンピューティング装置により実行されるプログラムモジュールなどのコンピュータ実行可能命令の一般的文脈においてさらに説明されることができる。一般に、プログラムモジュールは、特定タスクを実行し、又は特定抽象データタイプを実装するルーチン、プログラム、オブジェクト、コンポーネント、データ構造などを含む。本明細書に説明される実施形態は、1又は複数の通信ネットワークを通じてリンクされた1若しくは複数のリモートプロセッシング装置によって又は1若しくは複数の装置のクラウド内でタスクが実行される分散コンピューティング環境において行われてもよい。分散コンピューティング環境において、プログラムモジュールは、媒体記憶装置を含むローカルコンピュータ記憶媒体とリモートコンピュータ記憶媒体との双方に位置することができる。またさらに、前述された命令は、部分として又は全体として、ハードウェアロジック回路として実装されてよく、このハードウェアロジック回路は、プロセッサを含んでよく、あるいは含まなくてよい。

【 0 0 5 8 】

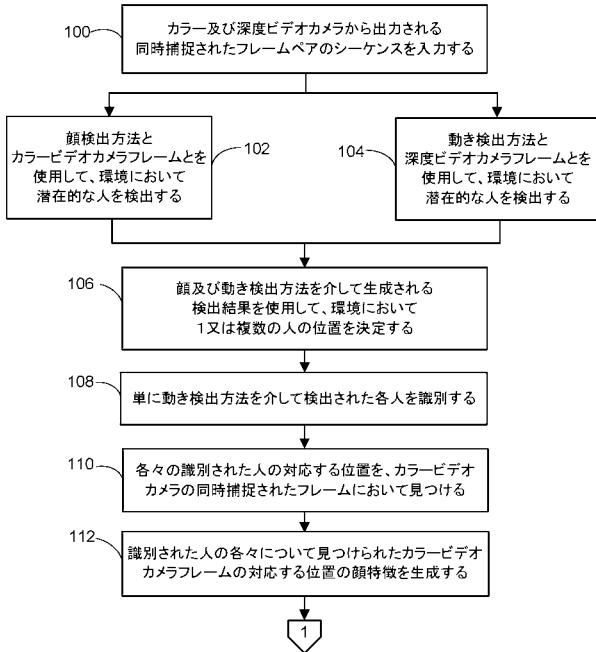
4 . 0 他の実施形態

前述の顔認識トレーニングデータベース生成手法実施形態の説明において、深度ビデオカメラなどのカメラから深度フレームを使用する、深度ビデオカメラと動き検出方法とが採用された。しかしながら、単にカラービデオカメラを用いて環境において人を検出することができる従来の動き検出方法もある。上記を仮定すると、別の実施形態において、深度ビデオカメラは取り除かれ、単にカラービデオカメラが使用されて環境において潜在的な人を検出する。したがって、これまでに説明されたプロセスは、カラービデオカメラから出力されたフレームのシーケンスが入力されるように変更されることになる。それから、これらの画像フレームは、環境において潜在的な人を検出する顔検出方法と共に、及び上記環境において潜在的な人をさらに検出する適切な動き検出方法と共に、使用される。同様にして、フレームの新しいシーケンスがこれまでに説明されたとおり採用されるとき、これらもまた単に、カラービデオカメラから出力されるフレームの新しいシーケンスとなる。

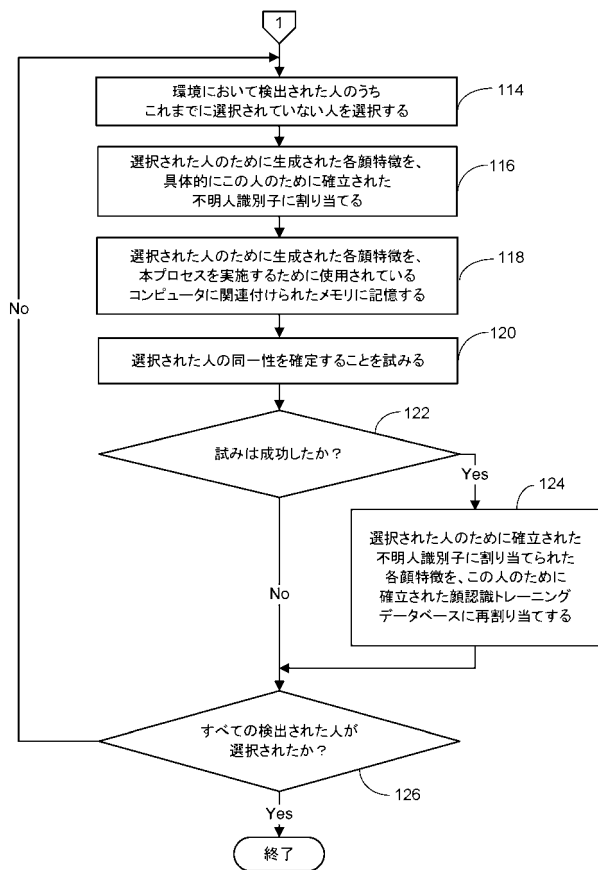
【 0 0 5 9 】

本説明にわたって前述された実施形態のうち任意のもの又はすべてを、所望される任意の組み合わせにおいて使用して、追加の混成の実施形態を形成することにさらに留意されたい。さらに、本願主題は構造的特性及び／又は方法論的動作に固有の言語で説明されているが、添付された請求項に定義される主題は、必ずしも上記で説明された特定の特性又は動作に限定されないことを理解されたい。むしろ、上記で説明された特定の特性及び動作は、請求項を実施する例示的形態として開示されている。

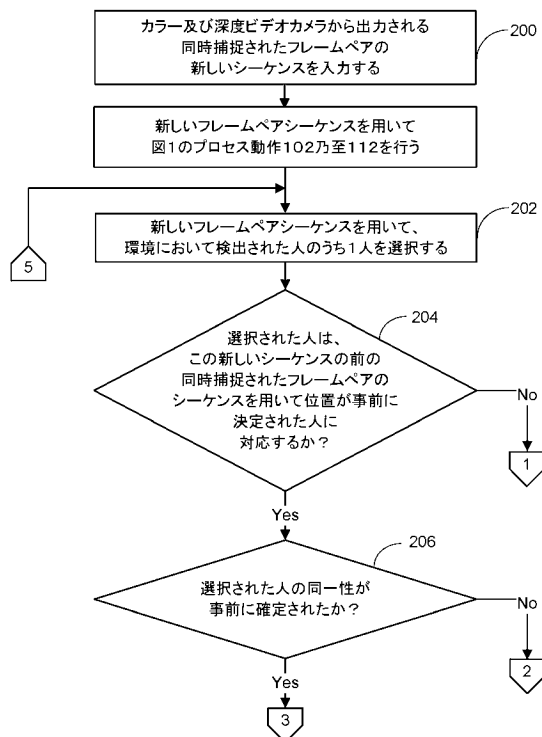
【図 1 A】



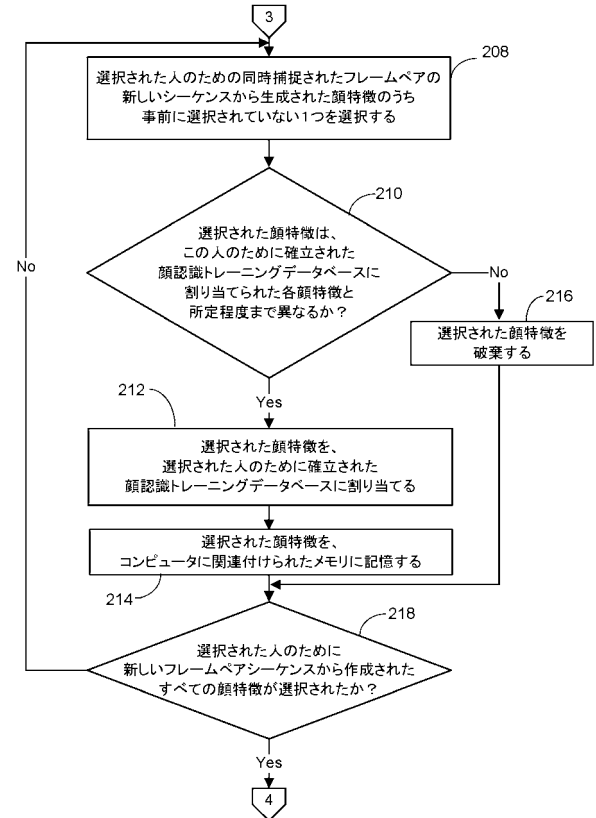
【図 1 B】



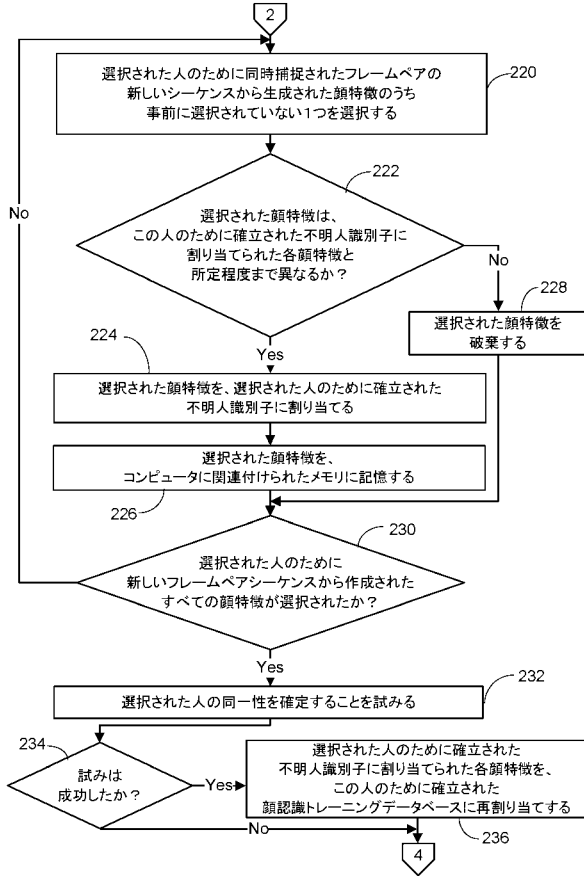
【図 2 A】



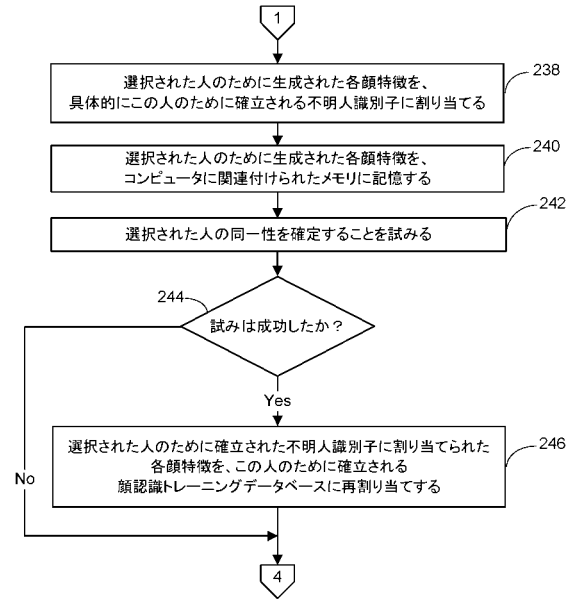
【図 2 B】



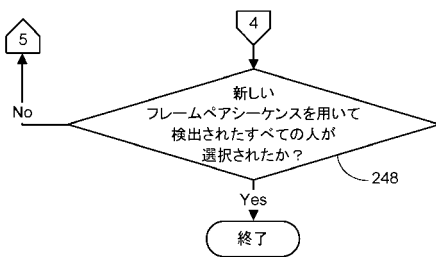
【図 2 C】



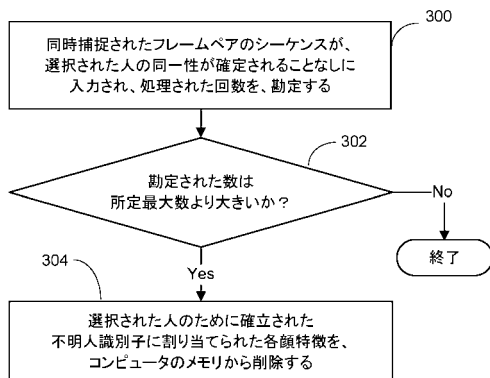
【図 2 D】



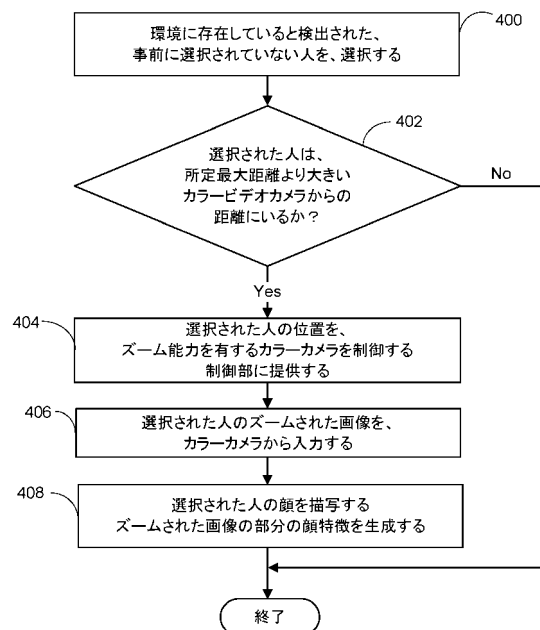
【図 2 E】



【図 3】

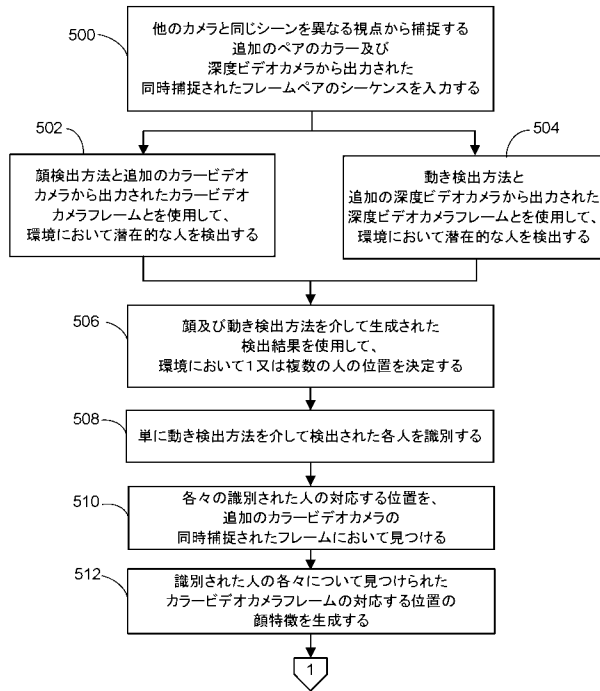


【図 4】

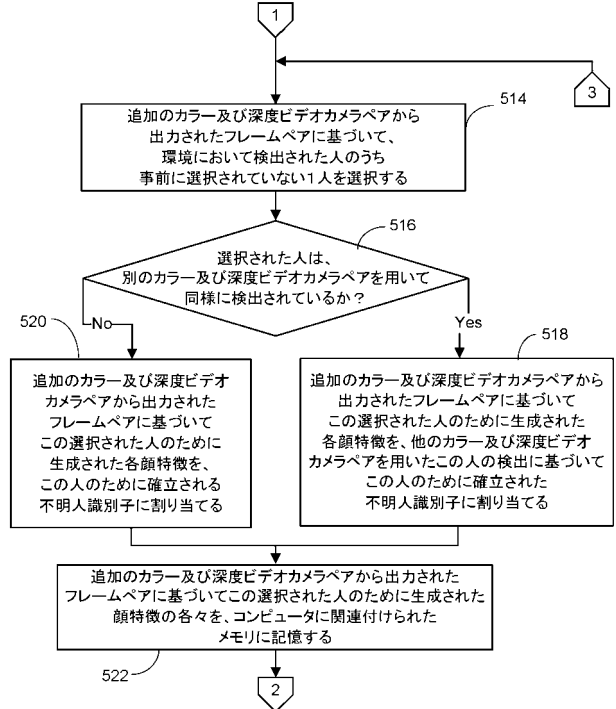




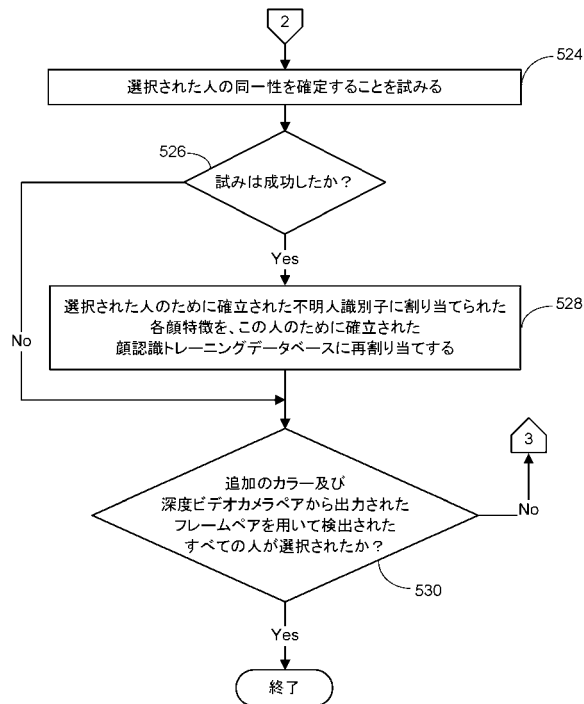
【図 5 A】



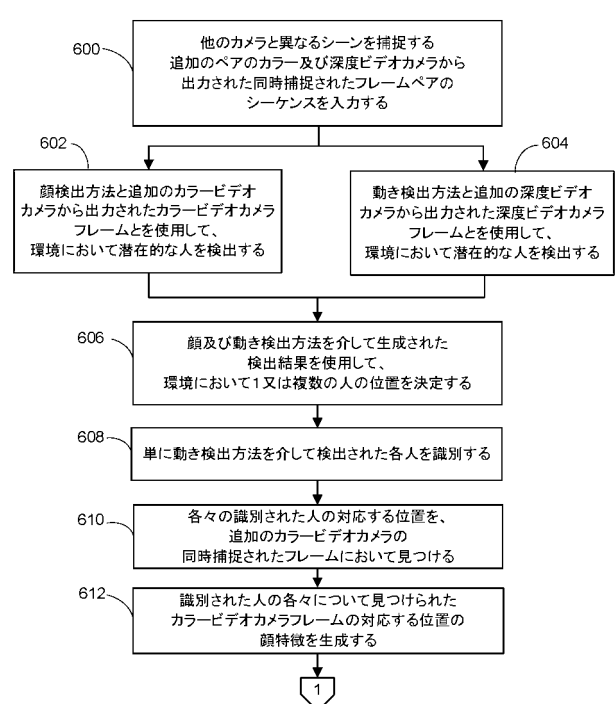
【図 5 B】



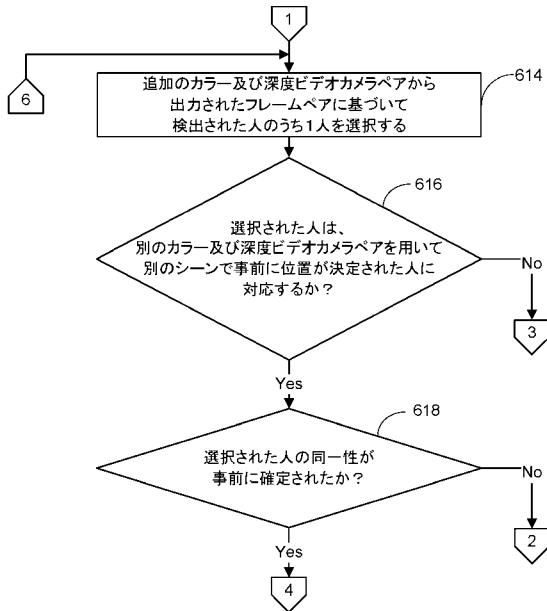
【図 5 C】



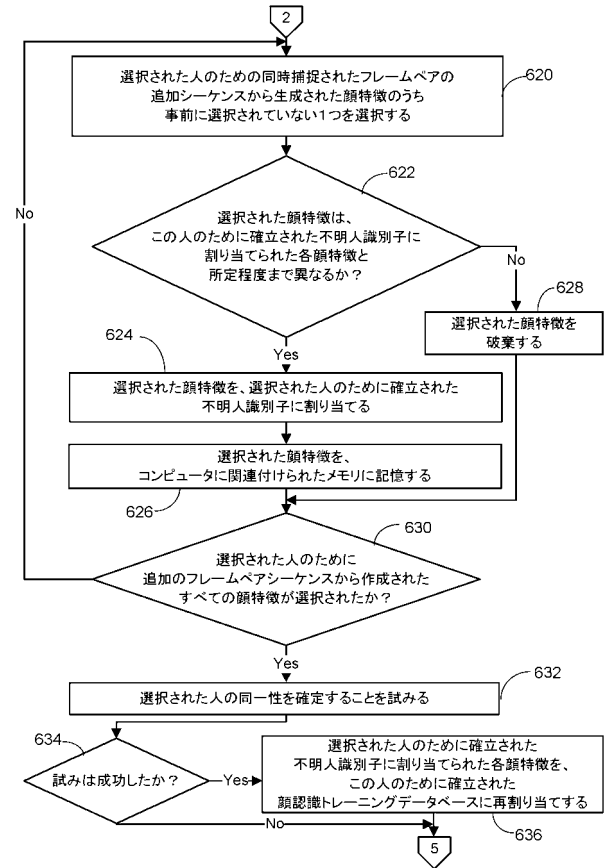
【図 6 A】



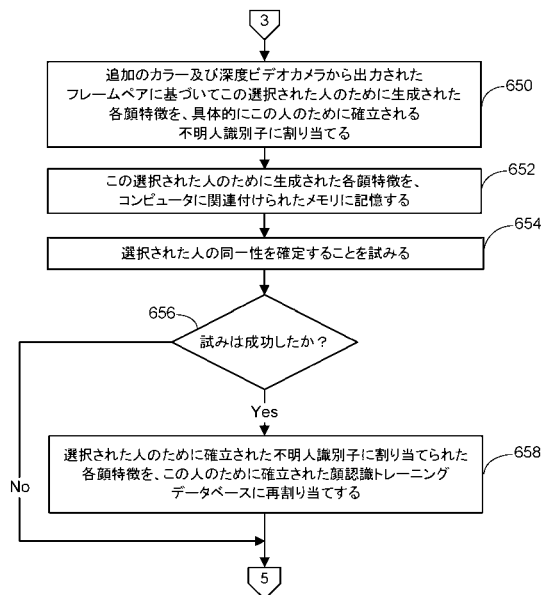
【図 6 B】



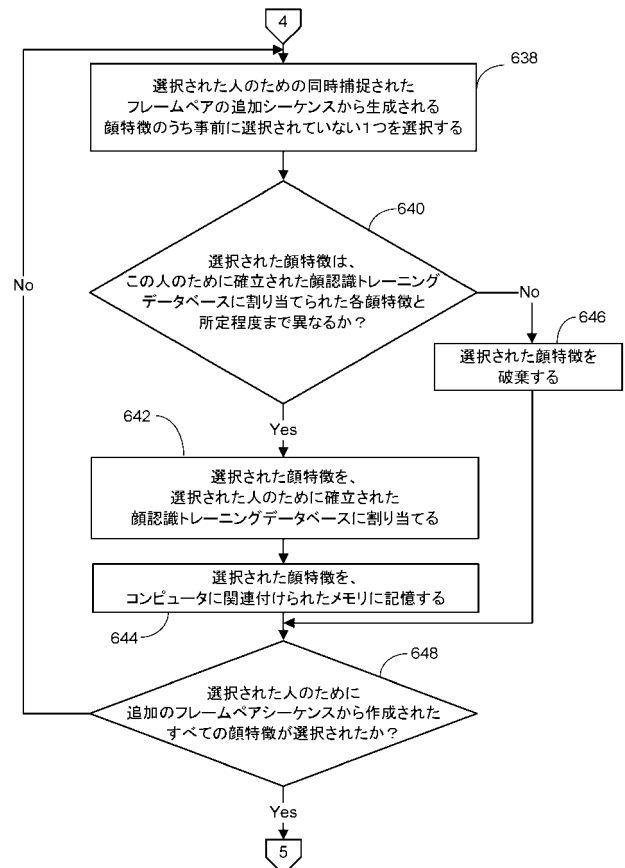
【図 6 C】



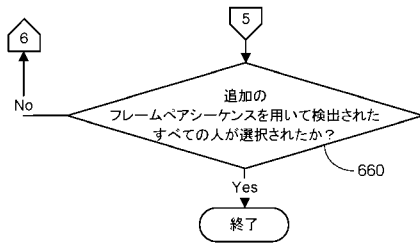
【図 6 D】



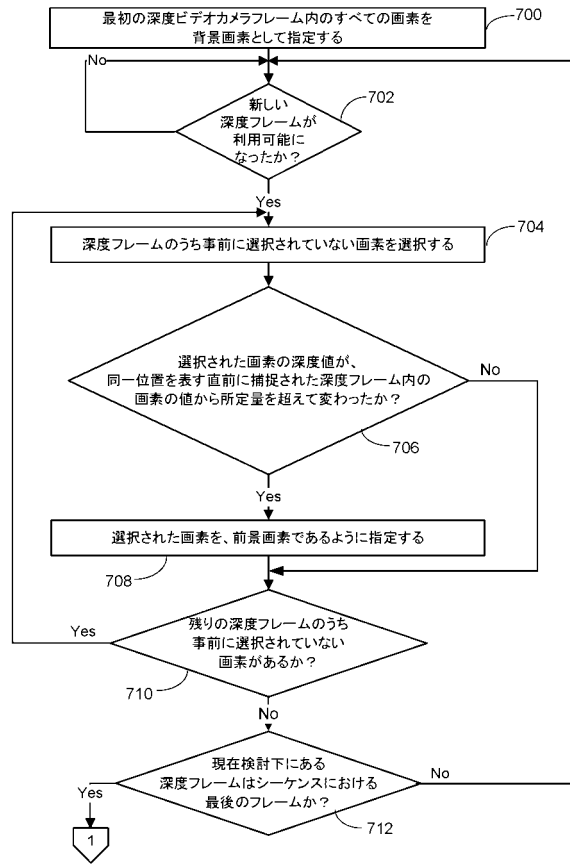
【図 6 E】



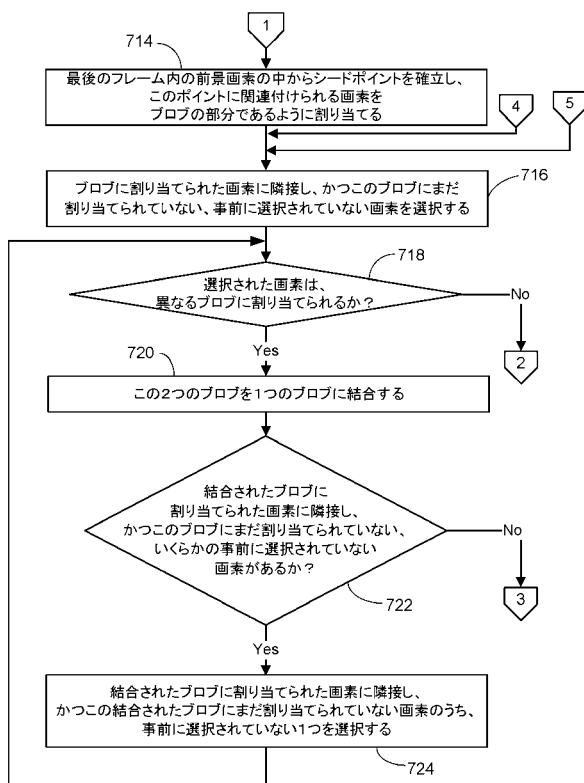
【図 6 F】



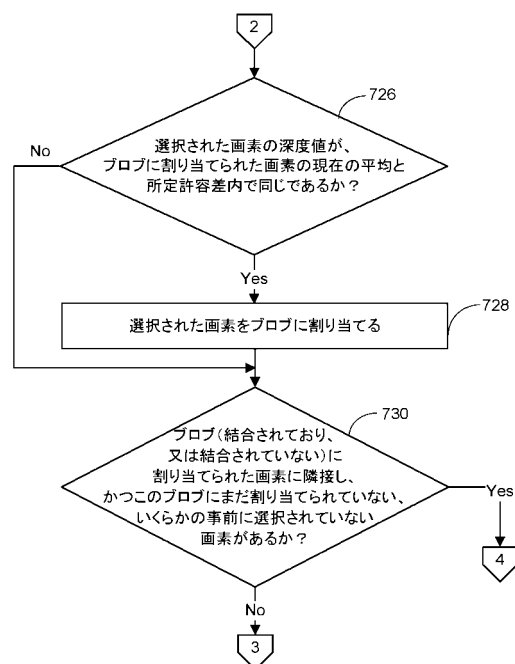
【図 7 A】



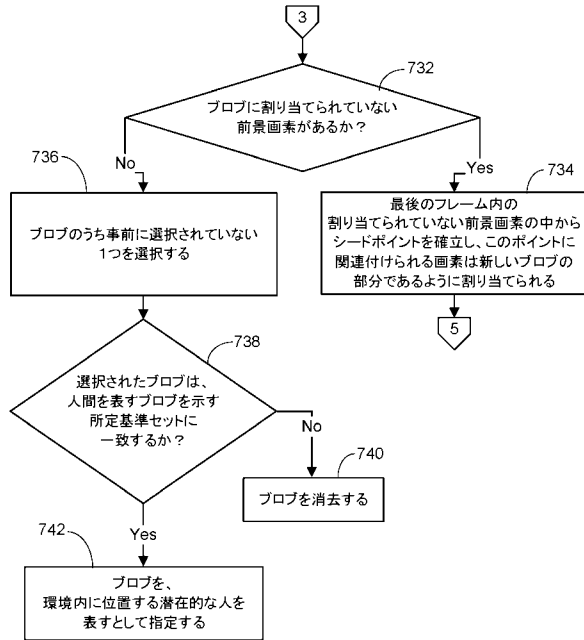
【図 7 B】



【図 7 C】



【図 7 D】



【図 8】

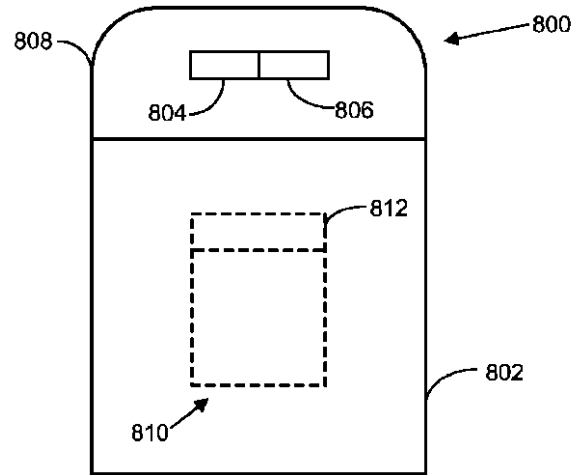
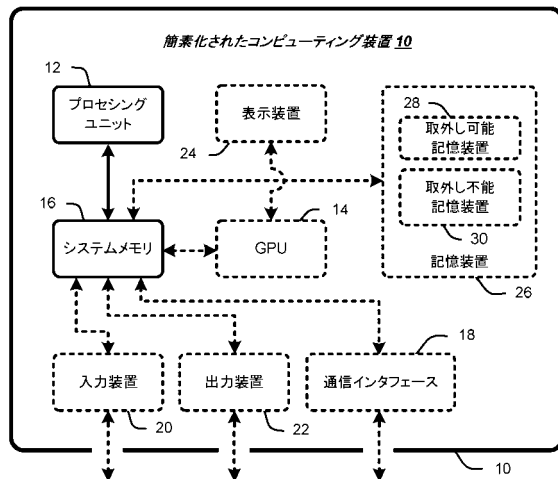


FIG. 8

【図 9】



## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2013/046447

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. G06K9/00 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
EPO-Internal		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2012/057665 A1 (ERICSSON TELEFON AB L M [SE]; FROEJDH PER [SE]; STROEM JACOB [SE]) 3 May 2012 (2012-05-03) page 1, lines 27-30 page 2, lines 3-11, 16-23 page 3, line 24 - page 4, line 11 page 10, line 15 - page 11, line 17 -----	1-10
Y	DAN XU ET AL: "Integrated approach of skin-color detection and depth information for hand and face localization", ROBOTICS AND BIOMIMETICS (ROBIO), 2011 IEEE INTERNATIONAL CONFERENCE ON, IEEE, 7 December 2011 (2011-12-07), pages 952-956, XP032165908, DOI: 10.1109/ROBIO.2011.6181410 ISBN: 978-1-4577-2136-6 Section I, III A and III C -----	1-10
-/-		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
11 September 2013		17/09/2013
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer
		Meurisse, Wim

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/046447

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DARRELL T ET AL: "Integrated person tracking using stereo, color, and pattern detection", INTERNET CITATION, [Online] XP002198613, Retrieved from the Internet: URL:http://www.ai.mit.edu/~trevor/papers/1998-021/TR-1998-021.pdf> [retrieved on 2002-05-10] Section 1, 4 first paragraph; figure 1 -----	1-10
A	WO 2012/071677 A1 (TECHNICOLOR CHINA TECHNOLOGY CO LTD [CN]; ZHOU JIE [CN]; CHENG PU [CN]) 7 June 2012 (2012-06-07) page 1, line 22 - page 3, line 10 page 13, lines 4-9 -----	1-10

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/US2013/046447

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2012057665	A1	03-05-2012	EP 2633474 A1 04-09-2013
			US 2013177219 A1 11-07-2013
			WO 2012057665 A1 03-05-2012
-----			
WO 2012071677	A1	07-06-2012	NONE
-----			

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC

(72)発明者 キッケリ, ハルシャヴァルダナ ナラヤナ

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト  
ウェイ マイクロソフト コーポレーション エルシーエー - インターナショナル パテンツ 内

(72)発明者 ケーニツヒ, ミヒャエル エフ.

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト  
ウェイ マイクロソフト コーポレーション エルシーエー - インターナショナル パテンツ 内

(72)発明者 コール, ジェフリー

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト  
ウェイ マイクロソフト コーポレーション エルシーエー - インターナショナル パテンツ 内

Fターム(参考) 5L096 AA02 AA06 BA18 CA04 FA66 FA69 HA03 JA03 JA09 JA11

KA04 KA13